Australian Government
Department of Defence
Defence Science and
Technology Organisation

A Review of Registration Capabilities in the Analyst's Detection Support System

Ronald Jones, David M. Booth, Peter G. Perry and Nicholas J. Redding

**Intelligence, Surveillance and Reconnaissance Division
Information Sciences Laboratory**

DSTO–TR–1632

ABSTRACT

This report presents a review and classification of image registration methods that are either currently available in the Analyst's Detection Support System (ADSS) or scheduled for implementation in ADSS in the near future. The aim of this report is to gain an overall understanding of our capabilities in the field of image registration, by identifying key techniques that we are using, highlighting instances where techniques could be reused to augment other methods, and identifying areas of methodology that need further development. In so doing, we aim to gain an understanding of where future work might best be directed in order to meet our current task goals in image registration.

Published by

DSTO Information Sciences Laboratory
PO Box 1500
Edinburgh, South Australia, Australia 5111

Telephone: (08) 8259 5555
Facsimile: (08) 8259 6567

AR No. 013-218
March, 2005

APPROVED FOR PUBLIC RELEASE

A Review of Registration Capabilities in the Analyst's Detection Support System

EXECUTIVE SUMMARY

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints or from different sensors. Image registration is a crucial step in many image processing and vision applications and is widely used in remote sensing, medical applications and computer vision; it is a broad field with a wide range of methodologies and techniques. This report presents a review of the registration methods that are either currently available in the Analyst's Detection Support System (ADSS) or those which we have plans to implement in ADSS in the near future. The aim of this report is to provide some understanding of how our capabilities in registration are related to one another and how they are placed with respect to the field as a whole. In particular, we seek to identify areas where we have strong capabilities and areas where our capabilities need to be improved. In so doing, we aim to gain an understanding of where future work might best be directed in order to meet our current task goals in image registration.

At present, most of the registration methods in ADSS are at various stages of completion, and a few are at the beginning stages of development, in particular those that deal with video image processing. We are currently well placed then to consider future directions before embarking on further development in video registration. We also find that there is room for growth in the area of feature-based registration methods. ADSS has extensive capabilities in feature detection, but the capabilities have not been applied directly to the registration problem. Finally, one of the key current task objectives is to perform real-time georeferencing of motion imagery with other forms of geo-imagery, *e.g.*, registering a video sequence from a flyover with aerial photography. This is essentially a scene-to-model registration problem and it is apparent that we currently have few direct capabilities within ADSS to perform this method. It would seem worthwhile then to devote further attention to this with a view to charting a way forward in our efforts in image registration.

Authors

Ronald Jones
Intelligence, Surveillance and Reconnaissance Division

Ronald Jones received a B.Sc. and Ph.D. in Physics from Victoria University of Wellington and Monash University, Melbourne, in 1990 and 1994, respectively. Since completing his PhD, he has worked as a Research Scientist at the CSIRO, Division of Mathematical and Information Sciences, and Proteome Systems Ltd, a BioTech company based in Sydney. He joined DSTO as a Senior Research Scientist in 2004. His research interests currently include linear and nonlinear image processing techniques.

David Booth
Intelligence, Surveillance and Reconnaissance Division

David Booth has a BSc(Hons)degree in Computer Science (1984), an MPhil in Remote Sensing (1986), and a PhD in Pattern Recognition (1991). He has been engaged in scientific research for The Ministry of Defence (UK) since 1986, mainly in the imagery exploitation field. During that time he led the ISTAR Imagery Exploitation Research Team at DERA Malvern for five years, and subsequent to that, the Image Understanding Project Team at DSTL Malvern. He is currently on long term detached duty at DSTO. He is a Chartered Engineer and Fellow of the British Computing Society

Peter G. Perry
SYDAC Pty Ltd

Peter Perry has an M.Sc in Computing from the University of Adelaide. He lectured in mathematics at Salisbury C.A.E. for 10 years before joining EMI Electronics Pty Ltd in 1982, staying with that company and its successors for 16 years and working mainly in the field of sonar processing. Since then he has had a number of contracting positions in commercial and defence software.

Nicholas J. Redding
Intelligence, Surveillance and Reconnaissance Division

Nicholas Redding received a B.E. and Ph.D. in electrical engineering all from the University of Queensland, Brisbane, in 1986 and 1991, respectively. In 1988 he received a Research Scientist Fellowship from the Australian Defence Science and Technology Organisation (DSTO) and then joined DSTO in Adelaide as a Research Scientist after completing his Ph.D. in artificial neural networks in 1991. In 1996 he was appointed as a Senior Research Scientist in the Microwave Radar Division (now Intelligence, Surveillance and Reconnaissance Division) of DSTO. Since joining DSTO he has applied image processing techniques to the automatic classification of ionospheric data, and more recently has researched target detection (both human and algorithmic) in synthetic aperture radar (SAR) imagery. He has recently returned from a one and a half year posting to the UK's Defence Evaluation and Research Agency where he continued the development of a suite of target detection algorithms for SAR imagery and researched new algorithms in SAR image forming using the circular Radon transform.

Contents

1 Introduction

This report presents a review of the registration methods that are either currently available in ADSS or those which we have plans to implement in ADSS in the near future. Image registration methods are widely used in remote sensing, medical applications and computer vision; there is a broad range of methodologies and techniques. One of the aims of this report is to provide some understanding of how our methods are placed with respect to the field and identify areas where our capabilities could be improved, in particular with regard to our current task goals. One way to achieve this is to show how our current capabilities are placed in proposed classification systems for registration methods, such as that presented in the review of Zitová and Flusser [28]. To begin with then, this report will deal primarily with methods by which registration methods may be classified, with reference to our particular methods.

Another aim of this report is to gain some understanding of how our registration methods are related to one another. In particular, what key techniques are being used and where techniques could be reused or leveraged to enhance or extend other methods. We are also interested in what key techniques we might be missing. To this end, we will be describing the generic processing steps involved in image registration, as per the review of Zitová and Flusser [28], and looking at how our methods fit within these processing steps. As it transpires, we seem to be favouring certain types of methods over others and there may be plenty of room for growth, in particular in the areas of local feature detection and matching.

This report will proceed as follows. In the following section, we introduce a classification system that may be used to classify the broad range of registration methods available; our current capabilities in ADSS will be classified according to this system. An alternative system which separates frame to frame methods from frame to reference methods is described in Section 2.5, as it is perhaps more pertinent to our interests in registration. We then proceed to describe the generic implementation steps that are used to implement any registration method in Section 3, again placing our capabilities with respect to the system. In Section 4, we will describe the registration methods in detail and highlighting areas where future work might be directed. Finally, a summary of the report is given in Section 5.

2 Classification of Registration Algorithms

Following the recent review of Zitova and Flusser [28], one way to classify the broad range of registration methods available is by the manner of image acquisition. Four main groups may be defined, as illustrated in Fig. 1 and defined below.

2.1 Multiview Analysis

For the purposes of multiview analysis, images of the same scene are acquired from different viewpoints. The aim here is to gain a larger 2D or 3D representation of the

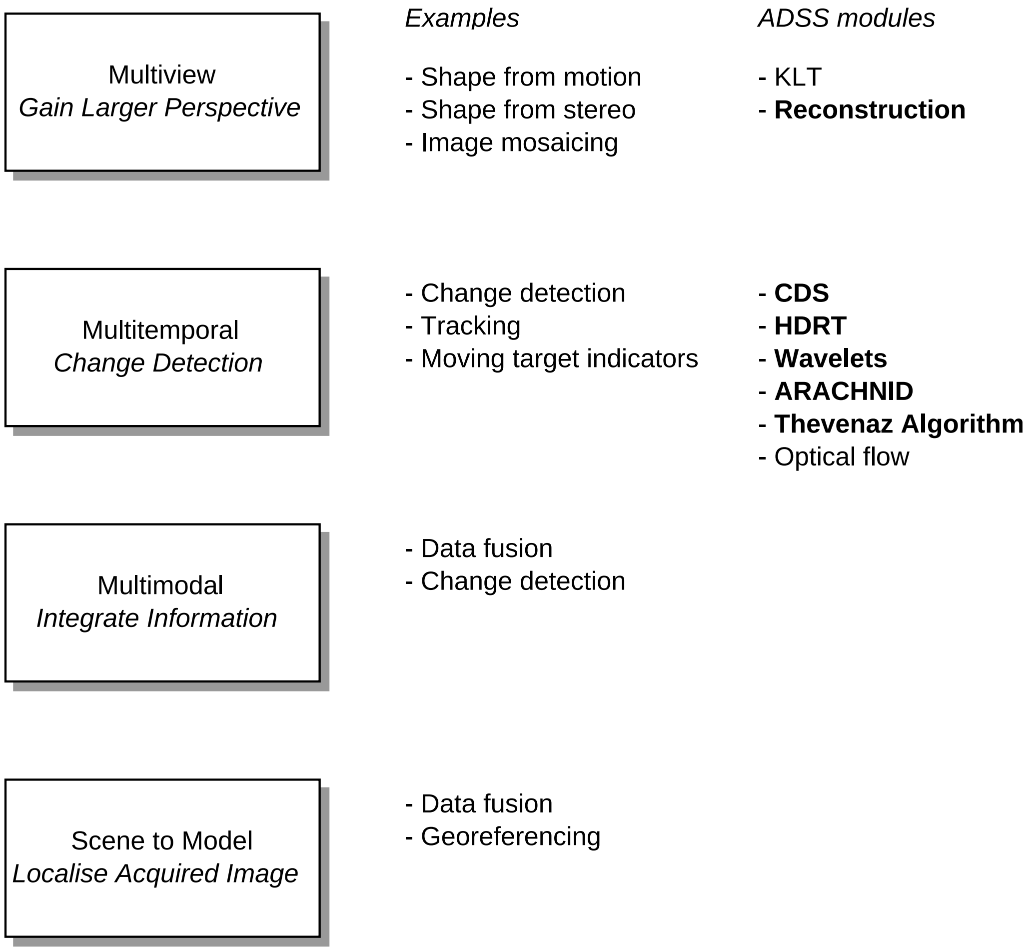

Figure 1: Registration Method Classification (I)

scene. Example applications include recovering 3D shape from stereo images or video and mosaicing of images of a surveyed area.

On the right side of the figure is shown registration methods that are currently implemented in ADSS [18] (in bold) or that we anticipate could soon be incorporated into ADSS. These modules will each be discussed in detail in Section 4; for now we will simply introduce and classify them into the most likely group.

- The "KLT" algorithm is the Kanade-Lucas-Tomasi feature tracker [21] and factorisation code [26] designed to track features points in video sequences and reconstruct 3D shape from motion.

- The "reconstruction" code, implemented in the ADSS modules `motion` and `matching` and related code, is an implementation of Phil Torr's structure from motion toolkit for Matlab and is based on (rather complex) techniques for 3D reconstruction from multiple view geometry [7].

2.2 Multitemporal Analysis

In multitemporal analysis, images of the same scene are acquired at different times, often on a regular basis, and possibly under different conditions. The aim here is to detect and evaluate change in the scene that occurs between image acquisitions. Example applications include automatic change detection, security monitoring and motion tracking.

This is where the bulk of the registration capabilities have been grouped; essentially because our problem domain currently focuses on change detection between pairs of images and on tracking in video.

- The ADSS Change Detection Subsystem (CDS) [19], also known as "JP 129", constitutes the model for image registration in ADSS. The method consists of three modular components: feature detection and matching using correlation in either the spatial or Fourier domain (as implemented by the modules `tie_points` and `tie_fft`), transform model estimation (`spline` module) and image resampling and transformation (`transform`).
- The HDRT method is an implementation of image registration using Hierarchical Discrete Radon Transforms [8, 17] and may be swapped into the CDS to perform the feature detection and matching step.
- Motion estimation and image registration using wavelets [12] is in the final stages of completion in ADSS; there are currently three modules implemented: `wavelets`, `motionField` and `motionResample`. Wavelets have seen broad application to motion estimation, change detection and shape reconstruction (*e.g.*, stereo reconstruction [13]).
- ARACHNID, or Automatic Registration and Change Detection, was developed by Dstl and QinetiQ Ltd and is a registration process based on correlation matching. An integral part of the methodology is to use one of a number of preprocessing steps to enhance features that are consistent over time. To this end, existing ADSS preprocessing modules can be utilised in a pipeline.
- The Thevénaz Algorithm, based on work by Thevénaz *et al* [23, 24], is a suite of code that is currently used by the tracking module `kalman_tracker`, the superresolution modules `multiframe` and `multi-tv` and the mosaicing module `mosaic0`. It provides the optimal affine transformation between a pair of image regions, based on a pyramidal decomposition.
- The optical flow based method, not currently implemented in ADSS, is based on work by Irani and Anandan [9] on moving object detection in 2D and 3D scenes. The method performs image registration by fitting affine transformations to a differential flow field. Matlab code to implement the method, supplied by Campbell-West and Miller [3], is currently under development.

2.3 Multimodal Analysis

In multimodal analysis, images of the same scene are acquired by different sensors. The aim is to integrate the information from different source streams to gain more complex and detailed scene representation.

One of the particular difficulties with multimodal registration is that actual image intensity values cannot be relied upon as the basis for image registration, as generally we cannot assume consistency between modes. Moreover, a certain amount of variation in scale must also be anticipated. In general, this excludes the use of area-based correlation methods in favour of methods based on local or higher level feature extraction and matching. Many registration methods can be applied to the multimodal case by using suitable preprocessing steps to extract or enhance image features consistent over the different modes. For example, the ARACHNID method can be used to register optical and infrared imagery using preprocessing steps to extract edges combined with positional information in the metadata. In the literature, methods based on *mutual information* [22] are leading edge for multimodal registration. They are based on measures of statistical dependency and have been used in medical image registration particularly.

Methods that are less suited to multimodal registration are those that require a high degree of overlap between frames and/or a simple model of image transformation, *e.g.*, those methods that are applied to unimodal video data such as the KLT method, wavelets, the Thevénaz Algorithm and the optical flow method.

2.4 Scene-to-Model Registration

In this group of methods, an image of a scene and a model of the scene are registered. The model could be *e.g.*, a computer representation of the scene, such as a map or a DEM in GIS. The aim is to localise the acquired image in the scene/model and/or compare them.

At present, there exists no registration method in ADSS that can perform a scene-to-model registration. However, as the stated primary objective of our efforts in shape from motion is to perform real-time georeferencing of motion imagery with other forms of geo-imagery, this is where we should be directing our future efforts in registration.

2.5 A Second Classification Method

A different classification method which perhaps more clearly characterises our two main interests in image registration, namely change detection and motion tracking, is shown in Fig. 2. Here the registration methods are classified into only two groups, again depending on the type of image acquisition, as described below. We we will see, they are special cases of the multiview, multitemporal and multimodal cases of the previous classification.

2.5.1 Frame to Frame Registration

In frame to frame registration, two frames in a video sequence are registered. Generally, we are dealing with a high volume of relatively small images (*e.g.*, 24 frames per second of 704×480 frames). We can usually assume a high degree of overlap between consecutive frames and so a fairly simple transform model, *e.g.*, a translation, similarity, affine or projective transformation. The registration method is implemented for each

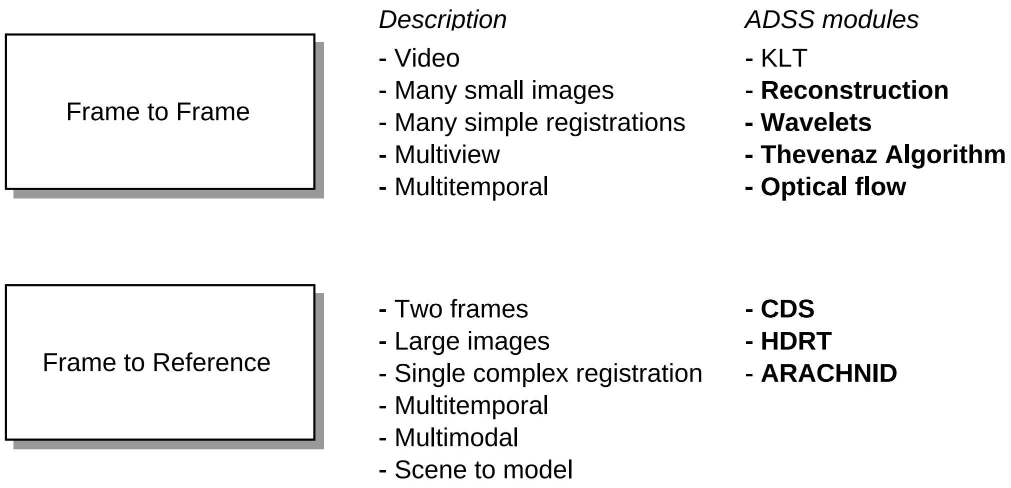

Figure 2: Registration Method Classification (II)

pair of consecutive frames in the video sequence and so should be fast but also accurate in order not to accumulate errors in *e.g.*, tracking applications. In terms of the previous classification system in Fig. 1, frame to frame registration is either multiview, in the case of a moving camera and a fixed scene (shape from motion), or multitemporal, in the case of a stationary camera and changing scene (moving target indicators). It can also be both multiview and multitemporal, in the case of a moving camera and moving scene, and this scenario represents some of the most challenging registration problems (*e.g.*, tracking moving targets from a moving platform, such as that performed by the module `kalman_tracker`).

As will be discussed in Section 3, a registration method may be divided into a series of distinct steps; broadly speaking: feature detection, feature matching, transform model estimation and image transformation. In the CDS registration method, these steps are implemented using distinct modules that are connected together within the ADSS framework to form the complete registration process. The power of such an approach is that it allows the swapping in or combining of alternative modules at any stage of the process in order to improve or refine the process for the particular application. As most frame to frame methods are still at an early stage of development, it may well be a good time to consider the broader picture of how best to implement these new methods within the ADSS paradigm of modular implementation and message passing.

As indicated in Fig. 2, the registration methods that could best be classified as frame to frame are the "KLT" and "registration" code, wavelets, Thevénaz Algorithm and optical flow.

2.5.2 Frame to Reference Registration

In frame to reference registration, two separate images, usually taken at quite different times, are registered. The images can be very large, *e.g.*, strip map SAR images with dimensions of many thousands of pixels. The registration of the two images is typically more complex and requires more time to implement. The transformation should accommodate a lower degree of image overlap and cater for local deformations. In terms of the previous classification system in Fig. 1, frame to reference registration is typically multi-temporal and finds application to change detection, mosaicing or data fusion. It may also be multimodal and/or be a case of scene-to-model registration. As indicated in Fig. 2, frame to reference registration is where the more mature registration methods in ADSS are grouped, in particular CDS, HDRT and ARACHNID. They are also the most modular and would seem to fit most comfortably within the ADSS modular development paradigm.

3 Implementation of Registration Methods

A typical registration method may generally be broken down into several distinct steps, as shown in Fig. 3. As will be discussed in more detail below, there are two main branches of registration methods. Feature-based methods first extract salient structures or features in the image and then implement a subsequent feature matching step to generate point correspondences, or "tie-points". In contrast, area-based methods use image areas or tiles to find the statistically best estimate of the translation vector. For the purposes of successive processing stages, this translation vector is considered to be a tie point with origin located at the centre of the image tile.

Both feature-based and area-based methods then feed into the subsequent transform model estimation and image resampling and transformation steps. It should be noted that although conceptually the series of steps used in image registration can be considered separately, in practice they are often merged together in the interests of speed, efficiency, or effectiveness. For example, the feature matching step may be combined with the transform model estimation step in order to generate feedback into the feature matching algorithm (as is the case for the Thevénaz Algorithm and the RANSAC algorithm [7]).

3.1 Feature-Based Methods

This approach is based on the extraction of salient structures or features in the image, *e.g.*, significant points, lines or regions. The resulting features are called *control points* (CPs). The CPs should be distinct, spread through the image and detectable in both images. There is wide array of literature on feature detection in images, ranging from edge and corner detectors, to line detection and region segmentation algorithms. ADSS provides a number of modules to perform feature detection, in particular the Plessey [6] module for corner detection (also known as the Harris corner detector) and various prescreeners, which may also be considered feature detectors. Once features are detected in the image pair, the features are matched to form point correspondences between the pair of images, usually following some underlying model for the image registration. Matching can be based on *e.g.*, the grey levels in the neighbourhood of the CPs (local correlation), the feature spatial distribution (binary and grey level shape characteristics), or the spatial relationship between CPs. These matches may then be used as input to the transform model estimation step.

Feature-based methods are typically applied when the local structure information is more important or reliable than the information carried by the specific intensities. Its

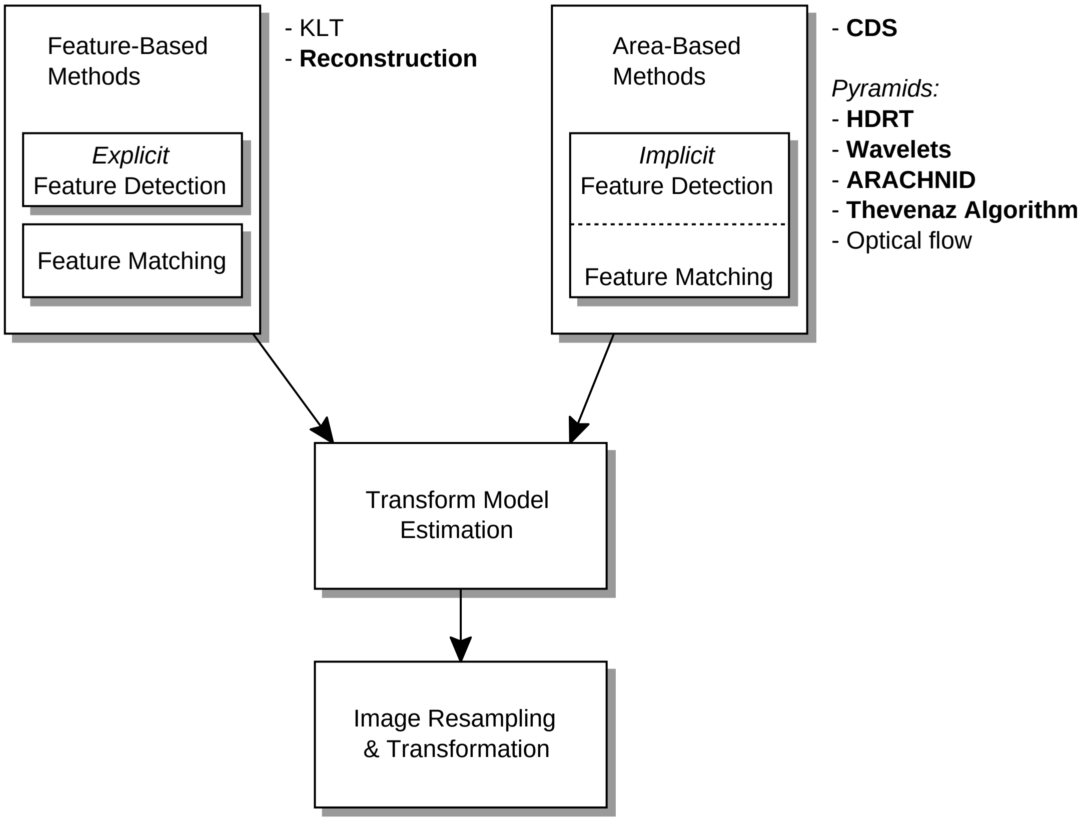

Figure 3: Registration Method Implementation

strength is that it can be used to match images of a completely different nature (*e.g.*, photos with maps) and can handle complex image distortions (*e.g.*, distortions of higher order than projective). A common drawback is that the features may be hard to detect or inconsistent in one or both of the images. To this end, it is important to use discriminative features and robust feature descriptors that are invariant to all assumed differences between the images. Moreover, the point correspondence problem can be difficult to solve, ill behaved and suffer from mismatches and crossovers. Methods using spatial relationships can be used if the detected features are ambiguous or if they are locally distorted; *e.g.*, graph matching, clustering or Chamfer matching [1].

It is interesting to note from Fig. 3 that, despite the fact that feature-based methods constitute an entire branch of registration methodology, none of the recognised registration methods in ADSS fall into this category. Only the "KLT" and "reconstruction" code, which are considered methods of generating shape from motion as opposed to image registration, fall into this category. This would suggest there is probably plenty of room for growth in ADSS in terms of developing further registration algorithms within this branch, complementing the strengths of the existing methods. Particular applications would be multimodal registration and registrations involving complex image distortions.

3.2 Area-Based Methods

This approach merges the feature detection and feature matching steps into a single step to produce statistically optimal translation vectors on the basis of matching (most often rectangular) image areas within the image. These are also called correlation or template matching methods. The matching process is typically a cross correlation (CC) technique that uses pixel intensities rather than local image structure to find the optimal translation vector between the two areas. Subpixel accuracy is possible using interpolation techniques and the frequency domain may be used to improve efficiency (in the case of larger windows) and remove frequency dependent noise.

Perhaps the greatest restriction on these methods is through the use of rectangular windows within which the correlation is carried out, as this restricts the local transform model to essentially a translation (although the global transform model may certainly be of higher order). Although CC methods can cope with some rotation and scaling, without appealing to generalisations of CC to more complex deformations, locally affine or projective transformations cannot easily be accommodated. Another drawback is that, because the approach is often based on a tiling strategy that is independent of image content, featureless regions can easily be matched together leading to high correlations and misleading results. The methods are also sensitive to noise in image intensities and are not well suited to multimodal registration without suitable preprocessing.

An important subgroup of the area-based methods are multiresolution registration methods based on coarse-to-fine strategies, or pyramids. The advantage with pyramidal methods is that matching with respect to the large scale image features is achieved at the coarsest resolution first, free from noisy perturbations at the local (finer) level. This robust result may then be used to guide matching at the next level of the pyramid, where the estimates are improved upon. The process continues down through the levels of the pyramid, thus achieving a progressively finer resolution of matching. At every level, pyramids significantly reduce the search space and thus save on the necessary computational time. The downside with using pyramids is that the strategy fails if a false match is identified at a coarser level in the pyramid; it is recommended that a backtracking or consistency check be incorporated into the algorithm.

Of the registration methods in ADSS that may be classified as area-based, all but the CDS method employ some kind of coarse-to-fine strategy: HDRT (employs image decimation), wavelets (inherently multiresolution), ARACHNID method and Thevénaz Algorithm (both employ pyramids), and optical flow (uses a three tier Gaussian pyramid).

3.3 Transform Model Estimation

The third step in the registration process is to use the feature correspondences to estimate the model transformation that maps one image to the other. Typically, an underlying model is assumed that is based on knowledge of the image acquisition and sensor characteristics. For example, a global model might be used such as simple translation, similarity, affine or projection. The correspondences are used to optimise the parameters associated with this transform. Depending on the choice of transform, there may be more correspondences than the minimum necessary to estimate the required transform. A least squares fit may be used so that the transform minimises the sum of the square errors at the feature locations.

The transform model may be classified into two types: global transformations, which use all correspondences for estimating one set of the mapping function's parameters valid for the entire image; and local transformations, which treat the image as a composition of patches and the function parameters depend on the location of their support in the image. Radial basis functions are an example of global mapping transforms that are able to handle locally varying geometric distortions. In particular, thin-plate splines are an example of radial basis functions that are currently implemented in the CDS registration method, by the module `spline`. Thin plate splines are known to give good results but the computations can be time consuming. Other splines that are used for image registration include B-splines and elastic body splines. Existing capabilities in ADSS for transform models will be discussed in more detail in the context of registration methods in Section 4.

3.4 Image resampling and transformation

In the final step of image registration, the sensed image (this could be considered the *frame* in frame to reference registration) is transformed by means of the mapping function into the reference image (or *reference* image). An appropriate interpolation technique is used to compute points falling between grid points (*i.e.* at non integer coordinates), *e.g.*, nearest neighbour or bilinear interpolation are often sufficient.

Image transformation can be done in either the forward or backward direction. In the forward direction, each pixel in the sensed image is directly transformed using the estimated mapping functions. However, due to rounding and discretisation errors, this can lead to holes and/or overlapping pixel values. For this reason, the backward direction is usually preferred. In this case, the inverse transform is computed and coordinates in the reference image are mapped to the sensed image domain, from which a pixel value is computed by interpolation. The ADSS module `transform`, which is part of the CDS method, employs a backward transformation and allows for a choice of four interpolation methods: nearest neighbour, bilinear, quadratic and least squares. The bilinear interpolant, though simple, offers a good trade off between computational cost and complexity. In contrast, nearest neighbour interpolation should be avoided in most cases because of artifacts in the sample image.

4 Current Registration Capabilities in ADSS

In this section, we will describe in more detail the registration methods that are currently available in ADSS or that we anticipate could soon be incorporated into ADSS. In particular, we report on each method in the context of classification and implementation methods, and highlight areas where future work might be done. We will look at those methods that have been classified as frame to reference first, followed by the frame to frame methods.

4.1 Change Detection Subsystem

The Change Detection Subsystem [19] (CDS), developed from the work by Nash [14] under Joint Project 129 (Airborne Surveillance for Land Operations), constitutes the current model for image registration, and is classified as a multitemporal, frame to reference, area-based method of registration. The method uses an area-based correlation technique to generate a sparse and evenly distributed set of likely feature correspondences that are subsequently modeled by a spline function and the imagery warped into alignment. Change detection may then be performed between the two images. The method consists of three modular components: feature detection and matching using correlation in either the spatial or Fourier domain, as implemented by the modules `tie_points` and `tie_fft`; transform model estimation via the `spline` module; and image resampling and transformation using the `transform` module. The modular approach allows for swapping in or combining of alternative modules at any stage of the process for the purposes of testing, improvement and refinement.

The `tie_points` module is an area-based feature matching process that attempts to identify feature matches in two images, covering approximately the same area, with approximately the same scale and orientation. The image is divided into tiles, where each tile is divided into sub-blocks. A correlation technique is used to find the best matching translation vector within each sub-block, and the sub-block with the highest correlation is retained provided that it agrees with sufficient other sub-blocks. The translation vector is considered to be a tie point with origin located at the centre of the image tile. Subpixel accuracy is possible using interpolation around the optimum point. There are two alternative interpolation methods: quadratic fit in both x and y directions; and least squares quadratic fit through 9 points. The method can also use positional information (*i.e.* image geocoding) as a starting point for registration; this means that smaller regions can be used provided that the unknown component of the translation is small.

The matching algorithm is subject to the following constraints: the size of the sub-block should be more than twice the registration error; there is minimal rotation between images; and the images have similar brightness and sufficient contrast. These are typical requirements on area-based correlation techniques and they are generally only suitable for simpler translations. Moreover, multimodal registration is not really possible with this method without suitable preprocessing due to the dependence on image intensities. The algorithm is therefore well suited to frame to frame registration, where there is high overlap between the frames to be registered and in particular minimal rotation and change of scale. At present however, the ADSS interface layer is written for image pairs and the test scripts are for large strip map images (8K $\times$ 13K pixels); these have been applied successfully to SAR, EO and IR imagery. It should be fairly straightforward for the code to be put into a library to be used more generally in frame to frame registration.

Correlation methods do not exploit structure in the image (*i.e.* image features) and are sensitive to errors from noise and different sensor types. Featureless regions can easily be matched together leading to high correlations and misleading results. The `tie_points` module is able to address this problem by using a threshold that specifies a minimum required contrast and a minimum number of non-zero pixels before a tie-point is generated.

The `tie_fft` module, which has not yet been documented, is also an area-based feature matching process but handles the correlation process in the frequency domain. In particular, following other Fourier methods, it estimates the translation between two areas by finding the corresponding phase shift in the frequency domain. Depending on window size this can be more efficient than using the spatial domain. It also has an algorithm for subpixel accuracy estimates and a quasi coarse-to-fine strategy based on truncation of the Fourier terms in the frequency domain. This is necessary because the method works best for small values of translations and a coarse-to-fine strategy allows a given translation to subtend larger regions at coarser levels. There are however outstanding implementation issues that have yet to be addressed, as noted in the code. In particular, a second pass over the data is still required to correct for phase wrapping and improve the translation estimate.

The `spline` module uses the tie-points generated by `tie_points` or `tie_fft` to produce a thin plate spline warping equation that maps the reference image to the sensed image. Thin plate splines are a global model of image registration, while allowing for local distortions. An option is available that allows the user to choose between an approximation model and an interpolation model for the thin plate spline. In particular, the approximation model specifies that the spline does not have to pass directly through the points given; the tolerance is given by a smoothing parameter. Due to the fact that ADSS processes images as streaming input data, a thin plate spline equation is generated for every set of three consecutive rows of tie-points. Typically then there may be many thin plate splines generated for any given image. The module works best if it receives all the tie-points at once and in this sense it is not readily parallelisable (although the speed of the algorithm does not appear to be an issue).

Thin plate splines are complex transform models that can handle both global mappings and local image distortions. In this sense they are best suited to the frame to reference registration problem that deal with large images with local distortions. If the registration area is small or if in particular we are registering frame to frame, the use of simpler transform models that have fewer degrees of freedom is more appropriate. For example, Caprari [4] has reported on image registration using a tiling strategy with small windows and local best-fit projective transforms. These transforms map a local square into a general quadrangle while preserving straight lines. The work applies mainly to wide angle images that require radiometric (intensity) registration. The work is essentially already present in ADSS through the implementation of the ARACHNID method, which is also based on optimal projective transforms, and would require a generalisation of `spline` or suitable replacement to actualise.

Finally, the `transform` module generates a registered version of the reference image with respect to the sensed image using a backward transformation. There are two modes of operation for this module. In tiled mode, the new image domain is divided into small tiles and the spline equation is used to find the corners of the tiles in the reference image. A projective transformation between the two is then calculated to map the reference to the new image. This process is fast provided tiles are not too small, however tiles that are too large start to introduce errors. The other mode of operation is to use the spline equations to map each individual point in the image; this is slower but is accurate. There are a number of options available within `transform` for pixel interpolation in the backward transformation, including nearest neighbour, bilinear, quadratic and least squares.

Based on the above observations, the following provides a summary of some of the areas where future effort on the CDS method could be directed:

- Application of `tie_points` code to frame to frame registration. Essentially, this could amount to putting the code into a library for use by other modules. However, it would also be good to make some decisions regarding the broader picture of how best to implement frame to frame registration within the ADSS paradigm of modular implementation and message passing, perhaps using `tie_points` as a test bed.
- Generalisation and consolidation of transform model estimation code. At present, only thin plate splines are implemented as a standalone method for transform model estimation (in the module `spline`), though there appears to be other examples of transform code throughout ADSS (*e.g.*, affine and projective transforms). Other transforms, *e.g.*, projective transforms, would be more appropriate for frame to frame registration. Work on this would tie in with broader decisions regarding the broader picture of how best to implement frame to frame registration.
- Completion of work on `tie_fft`. There are some unfinished elements of the code and there needs to be some documentation written for the module. In the longer term, there exist algorithms for the frequency domain for the implementation of correlation for rotated and scaled data that could be investigated.
- Modification of the algorithm to exploit phase information and as such have application to coherent change detection.

4.2 The Hierarchical Discrete Radon Transform Method

The Hierarchical Discrete Radon Transform (HDRT) method of registration [8, 17] is classified as multitemporal, frame to reference, and area-based. Radon transforms [2] have been used successfully to extract roads and faint trails in Synthetic Aperture Radar (SAR) imagery [5], as they are robust to background clutter and specular noise. The HDRT provides a hierarchy of Radon transforms, from the Radon transform of the entire image, right down to the Radon transform of single pixels in the image. The resulting HDRT structure is a coarse-to-fine pyramid that can be applied to hierarchically register images. Two separate ADSS modules are used in pipeline by the HDRT method: `hdrt`, which generates the actual pyramid of Radon transforms for a given image; and `registration`, which registers the HDRTs of two images and outputs the corresponding tie-points. These tie-points can then be fed to the `spline` and `transform` modules of the CDS method discussed previously.

The registration process starts at the coarsest level of the HDRT, where there are two Radon tiles that subtend the domain of the reference and sensed images. A correlation is then carried out between the two tiles, using a fast algorithm based on 1D convolution followed by backprojection. This provides a robust estimate of the global translation between the two images. It should be noted this method is mathematically equivalent to a standard 2D correlation in the spatial domain [16] (as used in the CDS method). That is, cross correlation in the spatial domain is equivalent to 1D correlation followed by back-projection in the Radon domain. In order to exploit the linear feature extraction capabilities of the Radon transform, a non-linear feature detection step is required. For example, thresholding of the HDRT in order to extract strong linear features. This would convert the current technique from an area-based method to a feature-based method and in so doing allow for multimodal image registation.

If there is a known global rotation between the two images, this is easily factored in with little additional computational cost using the Radon Shift Theorem. The translation estimation represents a weighted average of all translations taking place over the correlated tiles and may represent several different translations taking place within the correlated tiles. At present only the best estimate satisfying a specified threshold is used. Following the method of a pyramidal strategy, this coarse result is then used to guide matching at the next level of the pyramid, where the estimates are improved upon. The HDRT method implements a double overlapping (or four to one) tiling strategy that allows for a denser and more accurate matching of tiles. This guided process continues down through the pyramid to the desired level, thus achieving a progressively finer resolution of matching. A tie-point is then output for each match at this level.

One of the advantages of the pyramidal strategy is that it is able to register images that are separated by potentially large global translations. At the coarsest level, the search space is significantly reduced and thus it is possible to perform correlation over the entire image domain. This is in contrast to the CDS method for example, where correlation is done using subblocks at the original image resolution. Moreover, the coarse-to-fine strategy is able to gradually home in on local variations caused by terrain elevations and errors in global parameters. The use of the Radon transform has a smoothing affect on any specular noise in the image and this has particular application to registering SAR images. As the correlation is carried out on image features (straight lines), as opposed to image intensities in the spatial domain, it is less sensitive to local intensity variations, and could be well placed to handle multimodal registration.

However, the HDRT has the same restrictions as other area-based correlation methods: it is only well suited to predicting local translation estimates, as opposed to more complex local transformations. In particular, there should be minimal (unknown) rotation and scale variance between the areas to be matched and they should have similar brightness and sufficient contrast. This latter restriction can be mitigated by the use of a preprocessing step that normalises the image data to zero mean and unit variance (this step is unnecessary for complex imagery, as the mean of a complex image tends to zero). The method may not be well suited to frame to frame registration however, due to the high computational associated with constructing the HDRTs. Another potential downside of the method is that the pyramid strategy fails if a false match is identified at a coarser level in the pyramid. Backtracking or consistency checks should be incorporated into the algorithm.

Future work could be directed to the following areas:

- Introduction of a non-linear processing step to implement feature detection and allow for multi-modal image registration.

- Implementation of a backtracking or consistency check in order to cope with false matches at coarse levels. For example, an optimal graph search algorithm is presented in [10].

- Processing of multiple translation estimates. At present only the best estimate is handed down to the next level, when there may be several genuinely different translations taking place within the correlated tiles. This strategy could also be used to mitigate false matches.

- Extension of the algorithm to similarity matching. At present the matching algorithm handles a known global rotation using the Radon Shift Theorem. Preliminary discussions indicate this could be extended to unknown global rotations and, by extension, to unknown local rotations at finer levels of the pyramid. If unknown scaling could also be introduced, this would extend the current translation matching algorithm to a full similarity matching. Methods already exist in the frequency domain and these could be investigated [20].

- Implementation of interpolation for peak detection after correlation. Currently, no interpolation is implemented; it should be straightforward to apply existing interpolation methods in ADSS (*e.g.*, in the CDS method).

- Modification of the algorithm to exploit phase information and as such have application to coherent change detection.

4.3 ARACHNID

The ARACHNID (Automatic Registration and Change Detection) method of registration is classified as multitemporal, frame to reference, and area-based. The method, developed by Dstl and QinetiQ Ltd, finds the optimal projective transformation between an image pair using a correlation-based technique. The projective search space is explored from the outset using the whole image via a pyramidal decomposition of the image. After the optimal projective transformation is estimated at the coarsest level, the image is resampled before proceeding to the next level of the pyramid to refine the estimation. Once the images are registered at the finest level, change detection is then carried out.

The ARACHNID method is designed to work in concert with a number of image preprocessing algorithms. Their role is to make edges or high frequencies more prominent in some way, as it is the presence of boundaries between objects or ground cover that are usually consistent over time (as opposed to surface brightness which varies according to lighting, weather, season, *etc.*). Algorithms investigated include: intensity gradient, high pass filtering, local standard deviation, local entropy and local Eigen analysis. The latter three of these have been found to be the most effective. As the preprocessing steps can generally be pipelined from other existing ADSS modules they have not been carried over to ADSS from the original Dstl code.

The use of preprocessing steps and scale invariant transformations mean the ARACHNID method is well placed to handle multimodal image registration. In particular, it has the ability to register optical and infrared imagery, including the bi-modal case. This has been automated by using positional information held in the image metadata as the start conditions for registration. The projective transformation sets constraints that, in flat environments, are beneficial and are known to work well, *e.g.*, for frame to frame registration. However, for more general frame to reference registration, the technique needs to be extended to handle more varied terrain elevations where the transform would be inadequate.

At present, the ARACHNID module in ADSS is driven by a command of the form,

`(data cell (x y w h) (x0 y0 x1 y1 x2 y2 x3 y3 x4 y4)),`

where `(x, y)` is the top left corner of a cell of width `w` and height `h` centered on the detection in the main image, which approximately maps to the region with corners `(x0, y0)` ... `(x3, y3)` in the reference image. The corners are numbered clockwise from the top left corner and the reference coordinates may be fractional. For any given `data cell` command, the reference tile is aligned with the main tile and the projective transformation which achieves this is output as either a new `data cell` command, a tie-point, or a series of tie-points at the corners of the cell.

The following points indicate areas where future work might be directed:

- Application of method to frame to frame registration. At present, the ADSS interface layer handles only image pairs.
- Extension of the method to handle more complex global transforms other than projective transforms. This is expected to be carried out by Dstl some time in the future and hopefully will follow a modular design in the manner of *e.g.*, CDS.
- Documentation of the method. In particular, due to lack of documentation from QinetiQ, we do not know how the extensive set of sample parameters are used in the algorithm.
- A new module that uses pyramidal decomposition together with a simple search method to find a projective transform between two images is currently been investigated.

4.4 Complex Discrete Wavelet Transform

The Complex Discrete Wavelet Transform (CDWT) [11] is classified as a multitemporal, frame to frame, area-based transform. Wavelet decomposition has found application in stereo vision, shape from motion, motion estimation and image registration and could equally be classified as a frame to reference registration method. The CDWT provides a multiresolution decomposition of the image into a pyramid structure containing the high frequency image content at dyadically increasing scales in the image. The high frequency information is obtained at each level by applying a high pass filter with complex coefficients in both the vertical and horizontal directions separately. In the current implementation, this results in a set of six complex output images at each scale, corresponding to six different orientations in the spatial domain (paired in symmetry about the horizontal axis at angles 15, -15, 45, -45, 75 and -75 degrees). The residual low frequency information is passed on to the next level of decomposition in the pyramid. The strength of the wavelet representation is that it is able to characterise an image on the basis of generic image features, in particular arbitrarily oriented edges, at all scales within the image (from the local to the global scale).

In the CDWT method of registration [12], the six high frequency results at a given scale are compared using a similarity distance based on the square of the absolute value of pixel differences between the two images. An overall similarity distance is computed as the summation of the six individual results. Image matching using CDWTs is then an exercise in determining the translation that minimises the overall similarity distance at the given scale. At the coarsest scale, the process begins by finding the translation vector whose origin is at the centre of the reference image that has the minimum similarity distance. After a process of relaxation and smoothing, this translation is then bilinearly interpolated to the next finest level. The result is four translation vectors with origins at the centre of each quadrant in the reference image. The process continues until the desired fineness of scale is reached, forming what is know as a motion vector field (as produced by the module `motionField`). At this stage, the algorithm would be able to interface with the existing model of frame to reference registration, by outputting tie-points at the finest level. In such case, one would probably not generate the complete motion vector field, but stop part way down the pyramid and follow only a single path to the bottom level from each pixel at that level. The motion field has been used in ADSS to generate a panning video from two views of a scene (implemented by the module `motionResample`).

As has been mentioned above for other pyramidal schemes, hierarchical matching algorithms provide a means to reduce the complexity of matching over the entire image domain while keeping the same effective measurement range. This allows matching between images that do not have a high overlap; *i.e.* they are well suited to frame to reference registration. The disadvantage however is that it can impose vectors from coarse levels onto inappropriate regions of the finer levels and special strategies are required to recover from errors that are handed down from the top of the pyramid. Moreover, the registration method is again essentially correlation based and only appropriate for determining translations at the given scale, as opposed to more complex local distortions. The other potential disadvantage of pyramid strategies is the computational cost of producing a pyramid for each frame to be matched. This may detract from the application of CDWTs to video registration applications.

At this stage, the wavelet code in ADSS only works on two given image frames and so strictly speaking is not a complete frame to frame registration method. In order to generalise to a whole video sequence a generator script could be used to cycle over the frames. Thought is currently being given as to how this can be achieved more dynamically. In terms of future work, the following areas could be considered:

- Extension of the algorithm to frame to frame registration. At present, the modules in ADSS (`wavelets` and `motionField`) work with a pair of images but not a whole video sequence.

- Integration with current frame to reference model of registration, by using tie-points.

- Documentation of the method.

4.5 The Thevénaz Algorithm

The Thevénaz Algorithm [23, 24] is classified as a multitemporal, frame to frame, area-based transform. The algorithm provides the optimal affine transformation between a pair of image regions, based on a pyramidal image decomposition. More specifically, the pyramid is constructed using a cubic spline representation of the image and the optimisation of the affine transform is carried out using a modified Marquardt-Levenberg method. The registration method is closely related to the ARACHNID registration method, which seeks to find the optimal projective transformation between a pair of images using pyramidal decomposition (a projective transform has eight degrees of freedom; two more than an affine transform). The observations that apply to the ARACHNID method can also be applied here. In particular, the affine transformation sets constraints that are appropriate for frame to frame registration but often not frame to reference registration. It is therefore best suited to video applications and is currently used by a number of video processing modules in ADSS, including `kalman_tracker` (Kalman video tracking module), `multi-tv` and `multiframe` (video super resolution), and `mosaic0` (video mosaicing).

The algorithm is currently not implemented as a standalone registration module, but is available as a backend distribution that may be compiled for the given application. There are plans to rewrite the distribution and put it into a library, in part so that it can be distributed freely and in part because there is room for improvement. The key interface function is `regAffine`, which takes a pair of images, a region of interest and a set of tuning parameters, and returns the six parameters describing the optimal affine transform. Certain problems have been identified with the performance of `regAffine` however. In particular, it seems to be incorrectly matching images by skewing or shrinking the fragment onto the main image. It does not seem to give enough weighting to large unambiguous regions of the image that should be matched easily, despite experimentation with the tuning parameters and different masking regimes. Once the unexpected skewing begins, an image that is completely mismatched and heavily skewed is often produced a few frames further on. It also appears to have problems coping with noise; the tracking process is easily misled when matching against noisy pixels.

Some thoughts on where future work could be directed:

- Rewriting of distribution and putting it into a library.

- Addressing of some of the issues with the implementation to see if the problems can be fixed.

- Application of approach to frame to reference registration. ARACHNID employs a similar method and has been used successfully in frame to reference registration. The method may also be extended to handle more complex global transforms other than affine transforms.

- Supplementation of documentation. The code, which was downloaded from the website of Thevénaz [25], has minimal documentation unfortunately.

4.6 Optical Flow

The optical flow based method that we are considering implementing in ADSS is classified as a multitemporal, frame to frame and area-based registration method. It is based on work by Irani and Anandan [9] on moving object detection in 2D and 3D scenes and has been further evaluated in a report by Campbell-West and Miller [3]. The work is aimed at motion detection algorithms for affine sensor motions and is suited to frame to frame registration in video. The affine transformations are fitted to the differential flow field as derived from the methods of optical flow. The optical flow constraint requires that motion between frames be small, but this is extended through the use of a three-tier Gaussian pyramid decomposition. In particular, at the coarsest level of the pyramid, a shift of four pixels is represented by a one pixel shift satisfying the optical flow constraint. The method is similar to both the ARACHNID method and Thevénaz Algorithm, through the use of affine transform mappings and pyramidal decomposition.

The particular application of the method is to moving target detection in video sequences. The registration process is used to register consecutive frames before performing a local misalignment to identify moving targets. The scene is classified as a 2D scene when it can be approximated by a flat surface and a 3D scene when there are significant depth variations. The method provides a unified approach to handling moving-object detection in both 2D and 3D scenes, based on the stratification of the problem into scenarios which gradually increase in complexity. Currently, there is no C code or any modules in ADSS to perform the optical flow method of registration, although it is equivalent to the CDWT in its output and could be used as an alternative. However, Matlab code provided by Cambell-West and Miller [3] is in the process of being completed and tested, and this will enable a more thorough assessment of the algorithm. When implemented in ADSS, the optical flow method would provide a useful alternative to the module `kalman_tracker`, which implements tracking using the Thevénaz Algorithm in combination with a Kalman filter, and it may also be evaluated against the ARACHNID method. The disadvantage of the method is that, without further generalisation of the pyramidal approach, it seems to have limited application to frame to reference registration.

Some areas where future work could be directed are as follows:

- Obtain and test Matlab code implementation.
- Port code to ADSS, preferably within a modular framework of frame to frame registration.
- Evaluate against other methods, in particular the Thevénaz Algorithm and ARACHNID method.

4.7 KLT Feature Tracker

The combined KLT feature tracker [21] and factorisation method [26] is classified as a multiview, frame to frame, feature-based method of registration. Although the method provides a means of reconstructing shape from motion, as opposed to image registration per se, components of the process could be applied in methods that more directly deal with image registration, in particular those based on image features (as opposed to areas). The method relies on a high overlap between consecutive frames and it is based on image intensity comparisons; it is not well suited for frame to reference or multimodal registration.

The purpose of the KLT feature tracker is to identify and then track reliable features from frame to frame in a video sequence. In the context of frame to frame registration, these correspond to the two key mechanisms that underpin feature-based methods of registration; feature detection and feature matching (as illustrated in Fig. 3). The KLT method selects good features to track on the basis of optimising the overall tracking quality, as well as traditional measures of "interest" or "cornerness". Given the position of the feature in one frame, the position in the next frame is determined by finding the translation that minimises the dissimilarity over the (usually small) feature window. The quality of image features is monitored during tracking by using a measure of feature dissimilarity that quantifies the change of appearance of the feature between the first and the current frame. If the dissimilarity is too high, the feature is abandoned.

In the context of image registration, the KLT feature tracker generates a set of tie-points between consecutive frames in the image. Robustness can be enforced by requiring the feature be tracked over a given number of frames. The tie-points can then be used in the transform model estimation step of the registration process and the two frames registered. More specifically, an appropriate mapping function such as an affine or projective transform can be chosen for the assumed geometric deformation between frames. The associated parameters are then estimated by means of a least squares fit (in general we will have many more tie-points than we need to estimate the transform), so that the mapping function minimises the sum of square errors at the tie-points. In practice however, not all the tie-points will correspond to the background of the image, as features corresponding to moving objects will also be tracked. One of the key applications of frame to frame registration is to allow a simple pointwise comparison to expose independent object motion in the sequence. It is therefore desirable to avoid such tie-points where possible because they contribute to the registration of the moving objects as well as the background of the image. In such case, it would seem sensible to employ *e.g.*, the RANSAC algorithm [7], as discussed in Section 4.8, which provides a methodology for fitting transforms on the basis of an optimal subset of the observed tie-points.

Once feature tracks from the entire sequence of frames are extracted, the factorisation algorithm is then used to estimate the 3D positions of the feature points under the assumption of orthography, thus generating "shape from motion". The matrix of feature tracks is factorised into two separate matrices: The 3D structure of the feature points in the scene and the camera rotation parameters. In particular, the camera rotation parameters could have application to image registration because they specify how the camera is moving from frame to frame through the sequence. Given an assumed model of the scene (*e.g.*, a flat 2D scene), they may then be used to estimate the transform model for the purposes of registration. This would require further investigation.

The following areas could be considered for further work:

- Application of the KLT tracking method to a feature-based method of registration in video sequences, in particular use of the RANSAC algorithm.

- Investigation of the use of the factorisation method to determine transform models for frame to frame registration.

- Development of ADSS modules and documentation.

4.8 Reconstruction Code

The "reconstruction" code is an ADSS implementation of Phil Torr's structure from motion Matlab toolkit and is based on techniques for 3D reconstruction from multiple view geometry [7]. It is classified as a multiview, frame to frame, feature-based method of registration. There are currently two ADSS modules that use the code to generate shape from motion, `motion` and `matching`, but the code has not yet been applied directly to image registration.

The method is based on the detection of feature points in a pair of images using a Plessey corner detector [6] followed by correspondence point matcher. For any given feature in one image, the ideal corresponding point in the second image is that which has the maximum correlation in a local window. In contrast to the KLT tracker, the feature points are detected independently in each frame and matching is not restricted to consecutive frames. Rather, the user may specify a search distance that limits the length of the correspondence vector. The method would therefore be suited to both frame to frame and frame to reference matching. However, the use of the correlation window restricts the method to images of a similar scale and pixel intensity; it is not well suited to multimodal registration without the use of suitable preprocessing steps. In the context of image registration, the correspondence point matcher generates tie-points that can be used in a transform model estimation step.

For the purposes of generating structure from motion, the next step in the reconstruction process is to estimate the fundamental matrix F, which encapsulates the intrinsic projective geometry between the two views. It is independent of scene structure and depends only on the camera's internal parameters and relative pose. It is related to the camera rotation parameters generated by the factorisation method described in Section 4.7, and could potentially be applied to image registration. However, we are particularly interested in the RANSAC (RANdom SAmple Consensus) algorithm that has been used to estimate F on the basis of the tie-points [27]. The algorithm is quite general and has proven to be a very successful robust estimator that is able to cope with a large number of outliers. In particular, it should find ready application to the simpler problem of determining the transform model from a set of tie-points. The RANSAC algorithm may be summarised as follows [7]:

Objective

Robust fit of a model to a data set S that contains outliers

Algorithm

- Randomly select a sample of s data points from S and instantiate the model from this subset.

- Determine the set of data points S_i which are within a distance threshold t of the model. The S_i is the consensus set of the sample and defines the inliers of S.
- If the size of S_i (the number of inliers) is greater than some threshold T, re-estimate the model using all the points in S_i and terminate.
- If the size of S_i is less than T, select a new subset and repeat the above.
- After N trials the largest consensus set S_i is selected, and the model is re-estimated using all the points in the subset S_i.

The following areas could be considered for further work:

- Application of the reconstruction code to a feature-based method of registration, in particular using the RANSAC algorithm.
- Investigation of the use of the fundamental matrix F to determine transform models for registration.
- Development of further ADSS modules and documentation specifically for image registration.

5 Summary

At present, most of the registration methods in ADSS are at various stages of completion, and some of the methods are really at the very beginning stages, in particular those that deal with video image processing. At this time then, it would seem to make good sense to agree on and carry out any necessary additional work before implementing a study to compare the performance of the methods. A summary of some of suggestions for future work is given in Table 1.

In the broader context however, there are several conclusions we might draw from the study at this stage. The recognised registration methods we have implemented in ADSS so far all are based in what is known as area-based registration, where correlation type filters in rectangular windows are used to determine tie-points. As was illustrated in Fig. 3, area-based methods actually constitute only half of the recognised types of registration methods. The other half is based on (typically local) feature detection and feature matching. Although ADSS has extensive capabilities in feature detection, the capabilities have not been applied directly to the registration problem. Moreover, in order to further develop this branch of registration methods, it might be useful to give more thought to the characterisation of features in terms of their local spatial distribution (*e.g.*, moment analysis) and also the spatial relationships between features (*e.g.*, graphs). Other domains could also benefit from this work, *e.g.*, peak detection in Radon transforms could be improved using spatial characteristics (rather than just intensities). Work by Miller and Caprari [15] has also pointed out the usefulness of moment analysis in automatic target recognition.

The ADSS development paradigm could be described as one of designing separate modules that are linked by a message passing and generic image handling mechanism to

Table 1: Possibilities for future work on registration methods.

CDS	Apply to frame to frame registration Generalise transform model Complete work on `tie_fft`
HDRT	Backtracking and consistency checks Multiple translation estimates Similarity matching Interpolation of peak detection
Wavelets	Apply to frame to frame registration Integrate with CDS Documentation
ARACHNID	Apply to frame to frame registration Generalise transform model Documentation
Thevénaz Algorithm	Rewrite distribution and librarise Address implementation issues Apply to frame to reference registration Documentation
Optical flow	Obtain and test Matlab code Implement and document in ADSS
KLT	Apply to frame to frame registration Investigate application of factorisation Implement and document in ADSS
reconstruction	Apply to registration Investigate application of fundamental matrix Implement and document in ADSS

form complete systems. This "loose coupling" approach is powerful because it is highly extensible and allows rapid prototyping of imaging systems. For example, if code is available for an algorithm, it can simply be downloaded from the web, put into an ADSS module and then used directly within ADSS with other modules. To a certain degree then, we are interested in planning for modular implementations where possible. In particular, frame to frame registration methods appear at this time to be the least mature of our capabilities and we are currently well placed to consider design and implementation issues before embarking on further implementation.

The stated primary objective of our efforts in shape from motion is to perform real-time georeferencing of motion imagery with other forms of geo-imagery, *e.g.*, registering a video sequence from a flyover with aerial photography. This is essentially a scene-to-model registration problem and it is apparent that we currently have few direct capabilities within ADSS to perform this method. It would seem worthwhile to give this some further thought and discussion with a view to charting a way forward in our efforts in image registration.

References

1. Barrow I., Tenenbaum J. M., Bolles R. C., and Wolf H. C. Parametric Correspondence and Chamfer Matching: Two Techniques for Image Matching. Proceedings of the Fifth International Joint Conference on Artificial Intelligence, Cambridge, Massachusetts, 1977, pp. 659-663.

2. Bracewell R. N. *The Fourier Transform and its Applications, Third Edition.* McGraw Hill, 2000.

3. Campbell-West F. and Miller P. Evaluation of a Robust Least Squares Motion Detection Algorithm for Affine Sensor Motions. Submitted to IMVIP 2004.

4. Caprari R. S. Geometric and Radiometric Image Warping due to Displacement of Wide Field Optical Systems. Research Report DSTO–RR–0091, DSTO, 1996.

5. Cooke T. A Radon Transform Derivative Method for Faint Trail Detection in SAR Imagery. DICTA '99. Perth, Australia, 1999, pp. 31–34.

6. Harris C. G. and Stephens M. A Combined Corner and Edge Detector. Fourth Alvey Vision Conference, 1988, pp. 147–151.

7. Hartley R. and Zisserman A. Multiple View Geometry in Computer Vision. Second Edition, Cambridge Universtiy Press, March 2004.

8. Jones R., Redding N. J. and Payne M. Implementation of the Hierarchical Discrete Radon Transform with Application to Image Registration. Technical Report DSTO–TR–1599, DSTO Technical Report, 2004.

9. Irani M. and Anandan P. A Unified Approach to Moving Object Detection in 2D and 3D Scenes IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 6, June 1998.

10. Lew M. S. and Huang T. S. Optimal Multi-scale matching IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, June 1999, pp. 23–25.

11. Mallat S. G. A Theory of Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on Pattern Analysis Machine Intelligence, Vol. 11, No. 7, July 1989.

12. Margarey J. and Kingsbury N. Motion Estimation using a Complex-valued Wavelet Transform IEEE Transactions on Signal Processing, Vol. 46, No. 4, April 1998.

13. Margarey J. and Dick. A. Multiresolution Stereo Image Matching using Complex Wavelets. Proceedings of the 14th International Conference on Pattern Recognition, Vol. 1, August 1998, pg. 4.

14. Nash G. G. The Registration of SAR Images. Unpublished DSTO Technical Report, 1994.

15. Miller P. C. and Caprari R. S. Demonstration of Improved Automatic Target-Recognition Performance by Moment Analysis of Correlation Peaks. Applied Optics, Vol. 38, No. 8, 1999, pp. 1325–1331.

16. Payne T. M. Pre-formation SAR to SAR Image Registration. Proceedings IGARSS'01, Geoscience and Remote Sensing Symposium, Vol. 7, 2001, pp. 3033–3035.

17. Payne T. M., Margarey J. F. A. and Newsam G. N. Registration of Synthetic Aperture Radar Images Using a Multiresolution Radon Transform. Proceedings SPIE Vol. 3653, Visual Communications and Image Processing, 1999, pp. 394–405.

18. Redding N. J. Design of the Analysts' Detection Support System for Broad Area Aerial Surveillance. Technical Report DSTO–TR–0746, DSTO Technical Report, 1998.

19. Redding N. J., Blucher G., Booth D. M., Perry P. G., Jones R. and Hanan. R. The Analysts' Detection Support System: Architecture Design and Algorithms Part 2 - Algorithms. Technical Note DSTO–TN-0522, DSTO Technical Note, 2004.

20. Reddy B. S. and Chatterji B. N. An FFT-based Tehnique for translation, rotation and scale-invariant image registration. IEEE Transactions on Image Processing, Vol. 5, 1996, pp 1266-1271.

21. Shi J. and Tomasi C. Good Features to Track. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94), Seattle, June 1994, pp. 593-600.

22. Thevénaz P. and Unser M. A Pyramid Approach to Sub-pixel Image Fusion Based on Mutual Information. Proceedings of the IEEE International Conference on Image Processing ICIP'96, Laussane, Switzerland, 1996, pp. 265-268.

23. Thevénaz P., Ruttimann U. E. and Unser M. A Pyramid Approach to Subpixel Registration Based on Intensity. IEEE Transactions on Image Processing, Vol. 7, No. 1, Januray 1998.

24. Thevénaz P., Ruttimann U. E. and Unser M. Iterative Multi-scale Registration without Landmarks. Proceedings ICIP '95, Vol. III, Washington, USA, October 1995, pp. 228-231.

25. http://bigwww.epfl.ch/algorithms.html

26. Tomasi C. and Kanade T. Shape and Motion from Image Streams under Orthography: A Factorisation Method. International Journal of Computer Vision, Vol. 9, No. 2, 1992, pp. 137-154.

27. Torr P. H. S. and Murray D. W. The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix. International Journal of Computer Vision, Vol. 24, No. 3, 1997, pp. 271-300.

28. Zitová B. and Flusser J. Image Registration Methods: A Survey. Image and Vision Computing, Vol. 21, 2003, pp. 977-1000.

DISTRIBUTION LIST

A Review of Registration Capabilities in the Analyst's Detection Support System

Ronald Jones, David M. Booth, Peter G. Perry
and Nicholas J. Redding

	Number of Copies
DEFENCE ORGANISATION	
Task Sponsor	
DGICD	1
S&T Program	
Chief Defence Scientist FAS Science Policy AS Science Corporate Management Director General Science Policy Development	1
Counsellor, Defence Science, London	Doc Data Sheet
Counsellor, Defence Science, Washington	Doc Data Sheet
Scientific Adviser Joint	1
Navy Scientific Adviser	Doc Data Sheet
Scientific Adviser, Army	1
Air Force Scientific Adviser	1
Scientific Adviser to the DMO-ELL	Doc Data Sheet
Systems Sciences Laboratory	
Chief, Intelligence, Surveillance and Reconnaissance Division	1
Research Leader, Imagery Systems	1
Head, Image Analysis & Exploitation	1
Guy Blucher	1
Dr David Booth	1
Dr David Crisp	1
Dr Ronald Jones	1
David I. Kettler	1
Dr Nicholas J. Redding	4
Rodney Smith	1
Merrilyn Fiebig	1
Dr Mark Williams	1
Dr Vittala Shettigara	1
Ray Oermann	1
Bob Whatmough	1

Dr Nick J. S. Stacy	1
DSTO Library and Archives	
Library Edinburgh	2
Defence Archives	1
Capability Systems Division	
Director General Maritime Development	Doc Data Sheet
Director General Information Capability Development	Doc Data Sheet
Office of the Chief Information Officer	
Deputy Chief Information Officer	Doc Data Sheet
Director General Information Policy and Plans	Doc Data Sheet
AS Information Structures and Futures	Doc Data Sheet
AS Information Architecture and Management	Doc Data Sheet
Director General Australian Defence Information Office	Doc Data Sheet
Director General Australian Defence Simulation Office	Doc Data Sheet
Strategy Group	
Director General Military Strategy	Doc Data Sheet
Director General Preparedness	Doc Data Sheet
Navy	
SO (SCIENCE), COMAUSNAVSURFGRP, NSW	Doc Data Sheet
Director General Navy Capability, Performance and Plans, Navy Headquarters	Doc Data Sheet
Director General Navy Strategic Policy and Futures, Navy Headquarters	Doc Data Sheet
Army	
SO (Science), Deployable Joint Force Headquarters (DJFHQ)(L), Enoggera QLD	Doc Data Sheet
SO (Science), Land Headquarters (LHQ), Victoria Barracks, NSW	Doc Data Sheet
Intelligence Program	
DGSTA, Defence Intelligence Organisation	1
Manager, Information Centre, Defence Intelligence Organisation	1 (pdf format)
Assistant Secretary Corporate, Defence Imagery and Geospatial Organisation	1
Acquisition Program	
JP129 Program Manager	1
Defence Materiel Organisation	
Deputy CEO	Doc Data Sheet

Head Aerospace Systems Division	Doc Data Sheet
Head Electronic Systems Division	Doc Data Sheet
Head Maritime Systems Division	Doc Data Sheet
Head Land Systems Division	Doc Data Sheet
Defence Libraries	
Library Manager, DLS-Canberra	1
Library Manager, DLS-Sydney West	Doc Data Sheet
INTERNATIONAL DEFENCE INFORMATION CENTRES	
US Defense Technical Information Center	2
UK Defence Research Information Centre	2
Canada Defence Scientific Information Service	1
NZ Defence Information Centre	1
SPARES	
DSTO Edinburgh Library	5
Total number of copies:	43

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. CAVEAT/PRIVACY MARKING

2. TITLE	3. SECURITY CLASSIFICATION
A Review of Registration Capabilities in the Analyst's Detection Support System	Document (U) Title (U) Abstract (U)
4. AUTHORS	5. CORPORATE AUTHOR
Ronald Jones, David M. Booth, Peter G. Perry and Nicholas J. Redding	Information Sciences Laboratory PO Box 1500 Edinburgh, South Australia, Australia 5111

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO–TR–1632	013-218	Technical Report	March, 2005

8. FILE NUMBER	9. TASK NUMBER	10. SPONSOR	11. No OF PAGES	12. No OF REFS
2004/1084978	INT04/028	ASC(DIGO)	25	28

13. URL OF ELECTRONIC VERSION	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO–TR–1632.pdf	Chief, Intelligence, Surveillance and Reconnaissance Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved For Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DEFTEST DESCRIPTORS

Image registration
Detection
Algorithms
Evaluation

19. ABSTRACT

This report presents a review and classification of image registration methods that are either currently available in the Analyst's Detection Support System (ADSS) or scheduled for implementation in ADSS in the near future. The aim of this report is to gain an overall understanding of our capabilities in the field of image registration, by identifying key techniques that we are using, highlighting instances where techniques could be reused to augment other methods, and identifying areas of methodology that need further development. In so doing, we aim to gain an understanding of where future work might best be directed in order to meet our current task goals in image registration.